Eye Movement Features in the Depth and Spatial Perspective Perception of Paintings and Other Static Scenes

Marsel Fazlyyyakhmatov*, Adelina Mutagirova, Oleg Nedopekin and Vladimir Antipov

Kazan Federal University, Russia

Abstract: In this paper, using the example of two scenes, they showed that it is possible to perceive the depth and spatial perspective of painting images (and other static scenes) without the condition of binocular disparity (hereinafter referred to as the 3D phenomenon). To study the 3D phenomenon, eye movement is recorded. The students of the Institute of Physics at Kazan Federal University and an experienced researcher of the 3D phenomenon, the author and the developer of the working methodology took part in the surveys on eye movement record. To identify the perception of spatial perspective, 3D raster images are used mounted on the same static stimulus scenes (pictures). The first plot includes obtaining information on eye movement in the conditions of perception of two paintings with the elements of a monocular perspective of images. Histograms show that students perceive the perspective of the pictures displayed on the monitor screen. In two scenes presented, eye focusing occurs outside the plane of stimulus images. When students demonstrate a monitor screen with text and a white sheet of the histogram, they show that the focus planes are located between the screen and the students' eyes.

Keywords: Eye Movement, Binocular Eye Tracker, Disparity, Binocular Depth Perception.

INTRODUCTION

Backus (Backus *et al*., 2001) states that Stereoscopic depth perception is based on binocular disparities.

It should be noted that it was possible to "ignore the observation of a picture as a flat "object" at the level of mental perception of painting. All of its images are on the same plane and perceive the artist's world as "soaring" in three-dimensional space (Gregory, 1997; Liu, & Nijhuis, 2020). However, Gregory does not show the way to carry out this process.

As a rule, motion effects are applied to obtain three-dimensional perception.

Eye movement has been used in various studies since the 19th century. The reviews over 50 years (Schutz *et al*., 2011) reflect the interest and various fields of application, and eye movement features. The appearance of binocular eye trackers greatly simplified the methodology for studying the processes of eye movement (Barabanschikov & Zhegallo, 2013; Zhegallo & Marmalyuk, 2015). For example, it is used to study binocular coordination (Bucci *et al*., 2008), the duration of fixations (Kliegl *et al*., 2006; Gómez-Huélamo *et al*., 2020).

The works (Angelaki & Hess, 2005; Iijima *et al*., 2012) showed, that a sensation of depth can arise from two-dimensional (2D) movies without any stereoscopic depth cue. Depth perception in three-dimensional (3D) space depends on the stability of stereoscopic gaze by vergence – coordinated movement of the two eyes in opposite directions – compensating the misalignment of the retinal images from the two eyes (i.e. binocular disparity) (Cazzato *et al*., 2020).

In our work, we use a binocular IT tracker to record eye movements and show the ability to perceive paintings (and other static scenes) with a spatial perspective. We continue to publish materials presented in (Antipov & Zhegallo, 2014; Antipov *et al*., 2018). We also use the graphical and mathematical method to record the 3D phenomenon described in (Antipov & Fazlyyyakhmatov, 2018). Note that the ability to perceive the depth of 2D image color palette was shown by us in the abovementioned publications. At the same time, it was noted in (Antipov & Fazlyyyakhmatov, 2018) that the use of eye movement in the study of the 3D phenomenon was mentioned in the section of the book (Barabanschikov & Zhegallo, 2013) as the "latest research on oculomotor activity". Some of our results were noted in (Yao *et al*., 2018; Abromavicius *et al*., 2018).

METHODS

The research methodology includes the selection of a stimulus 2D image creating the mentioned three-dimensional attributes (3D phenomenon). The choice of stimulus images is carried out by an experienced researcher (An-2), who has been studying the 3D phenomenon for a long time and perceives any planar images with the effects of depth, volume, and spatial perspective. For plot 1, we further develop a 3D raster

*Address correspondence to this author at the Kazan Federal University, Russia; E-mail: f_publication14@mail.ru

E-ISSN: 1929-4409/20

image containing similar depth effects that An-2 observed on a planar stimulus material. (Note that bitmaps are not used in scene 2 and 3). At the next stage, a stimulus 2D image is displayed on a monitor screen and the EEyTribe (Denmark) portable binocular eye tracker (recording frequency 60 Hz) is used to register eye movements with the determination of the X-coordinate arrays for the bi fixation points of the right (R) and left (L) eye. Eye movement recording time makes Δt = 60 s. The difference ΔX = X (R) - X (L) is found from the data arrays. Then the difference histograms are developed. The contour of the difference histograms (i.e., the probability density function) identifies the spatial location of the focus planes of the stimulus image relative to the monitor screen. Next, a 3D raster image is installed in front of the monitor screen and the eye movement is recorded again. The coincidence of the probability density functions obtained from the stimulus 2D image and the probability density functions from its 3D raster analogue allows us to conclude that the subject perceives the effects of the 3D phenomenon. The principle of calculating the location of focusing planes is shown in (Antipov & Zhegallo, 2014; Antipov *et al*., 2018).

A raster 3D stereo image is performed according to the technology providing its surround perception by the observer. Stereoscopic projections are used for its development. They are encoded with special software. Then the encoded image is applied to the paper and glued to the flat side of the plastic raster. A raster is a plate with a set of cylindrical lenses on one side. An adhesive layer is applied on the other, flat side of the plate. It is covered with a protective film. The coding creates vertical stripes on the plane of the paper sheet, the periodicity of which coincides with the periodicity of cylindrical lenses. When a coded paper print is connected to the adhesive base of plastic cylindrical lenses, an image intended for it is sent to each eye. In other words, the "geometry" of the raster ensures the refraction of the light flux so that each eye sees only the image intended for it. Note that to observe the depth of a 3D raster image, it is necessary to combine the encoded image on a paper printout with the flat side of the adhesive base of cylindrical lenses.

The methodology of a student interview in the field of analysis of paintings as stimulus images includes the acquisition of three difference histograms. The first is the histogram of a directly stimulated 2D image perception. The second histogram shows eye movement in the perception of coded paper projection. The third histogram captures the eye movement during a raster 3D image perception.

RESULTS AND DISCUSSION

Plot I

Figure **1** shows two stimulus images: a watercolor illustration "Kazan University" and the painting "The Seine near Giverny" (1888) by Oscar Claude Monet (https://www.wikiart.org/en/claude-monet/the-seine-near-giverny-1). The survey involved eleven 19-20 year-old students (4 young men), and An-2. Both images have monocular signs of perspective. When they consider stimulus images of the presented drawing format, all students claimed that they perceive the perspective of the images. This is a subjective perception.

Note that the location of the difference histograms in the region of positive ΔX values refers to the condition for perceiving the perspectives of stimulus images in Figure **1**. Histograms characterize the objective perception of depth and perspective.

a) b)

Figure 1: Stimulus images: **a**) watercolor illustration “Kazan University”, **b**) painting “The Seine near Giverny” (1888) by Oscar Claude Monet.

Stimulus Image – the Watercolor "Kazan University"

All students claim to observe the depth and perspective of the images.

- 8 (≈73%) students perceive the depth and perspective of the image behind the image display plane (Figure **2a-2c**). There are 2 young men among them. The histograms of all image variants are shifted to the region of positive ΔX values. The probabilistic values of the difference ΔX are plotted on the vertical scale of the figures with histograms, and conditional values ΔX are plotted on the horizontal scale. It does not matter that they are conditional. An important factor is only the difference sign. The position of the histograms, as on Figure **2a** and **2c**, indicates that the stimulus image is perceived behind the plane of the monitor screen. For three students, the difference histogram is shifted to the region of negative ΔX values, or the maximum falls to zero values (Figure **2b**).

The first information on the perception of a paper printout depth of an encoded image is given in (Antipov & Fazlyyyakhmatov, 2018).

Figure **2**: vertical rows are the difference histograms of two bachelors (a, b) and An-2 (c). Horizontal rows of figures: the top row is the perception of a 2D image; medium row is an encoded image; bottom row - the histograms of raster projection.

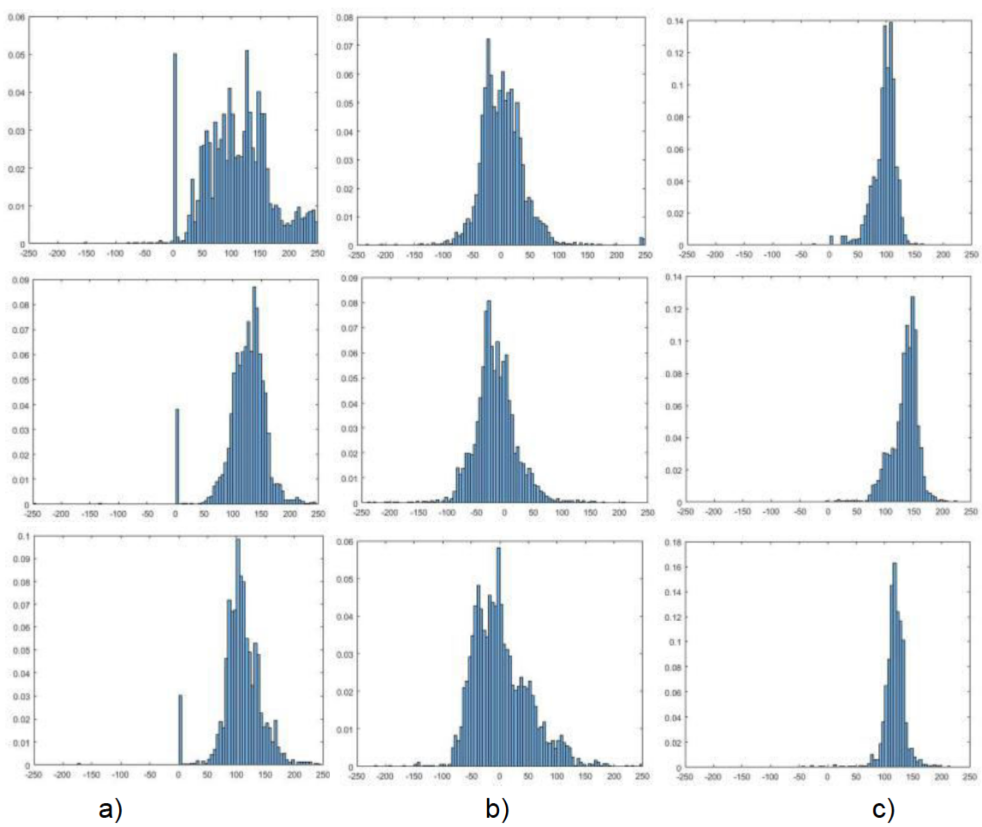
- for 6 (≈55%) students, the difference histograms during the perception of the 3D phenomenon, the encoded image (as for An-2) are located in the range of the raster image histogram (Figure **2a**, **2c**).

- for 5 (≈45%) students there is no coincidence of overlapping of all ΔX range record options.

The width of the difference histogram contour is determined by the individual characteristics of the subjects. Single vertical lines on the histograms show that there are also planar perceptual variations. The obtained histograms show that students perceive the depth of coded projections.

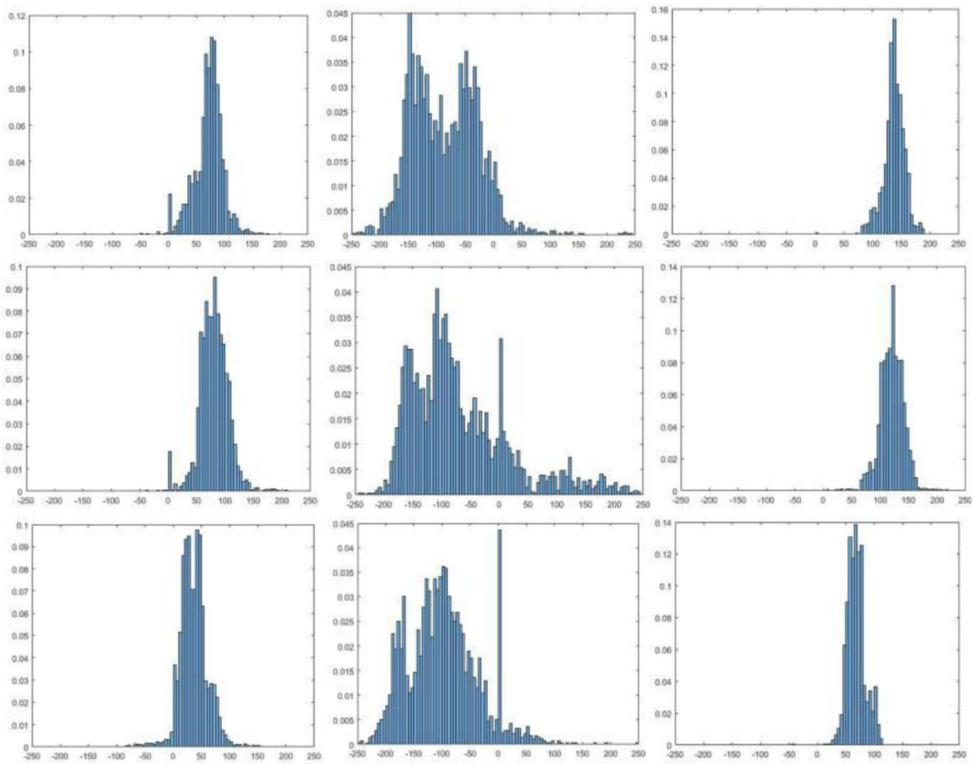
Stimulus Image - "The Seine near Giverny"

Figure **3** shows the difference histograms between two bachelors and An-2.



a) b) c)

Figure 2: Histograms in the perception of watercolors "Kazan University".



a) b) c)

Figure 3: Histograms during the perception of "The Seine near Giverny" picture.

Figure **3**: Vertical rows - difference histograms a), b) of two bachelors; c) An-2. Horizontal rows: upper - the histograms during the perception of 2D images; medium - on encoded images; lower - during a raster perception.

- 4 students (≈36%) perceive depth, the perspective of the image behind the monitor plane (Figure **3a**) as well as An-2 (Figure **3c**).
- for 3 students (≈27%), the maximum of the difference histogram contour falls into the region of zero readings ΔX.
- 1 student (≈9%) perceives the depth between the monitor screen and the eyes (Figure **3b**).
- for other students (≈27%), perception occurs either in the region of zero ΔX values, or between the eyes and the screen.
- the coincidence of the ranges of all three stimulus projections in front of the monitor plane was obtained for two students (≈18%).

Earlier, some histograms obtained for Figure **1b** were presented at the conference (Antipov *et al*., 2019).

Plot "II"

The use of binocular eye tracker allowed us to check the subjectively observed phenomenon of depth during the perception of the text on the monitor screen. Earlier, we suggested that when they type on the monitor screen, it may be separated from the white background of the sheet and long-term work behind the monitor screen contributes to the development of cognitive (and three-dimensional) vision skills (Minzaripov *et al*., 2009).

The survey involved 16 19-20 year-old students (8 young men), and An-2. 13 of them (≈81%) perceive text on the background of a white sheet between the monitor and the eyes.

In our works on the study of cortical activity under the conditions of 3D phenomenon perception (Fazlyyyakhmatov *et al*., 2018), the procedure includes EEG brain activity record when a white sheet is placed in the field of view. The eye movement of the student “k” under such conditions is shown on Figure **4c**, An-2 - Figure **4d**.

The comparison of the difference histograms of “k” and An-2 suggests that “k” perceives planar images at the level of 3D phenomenon.

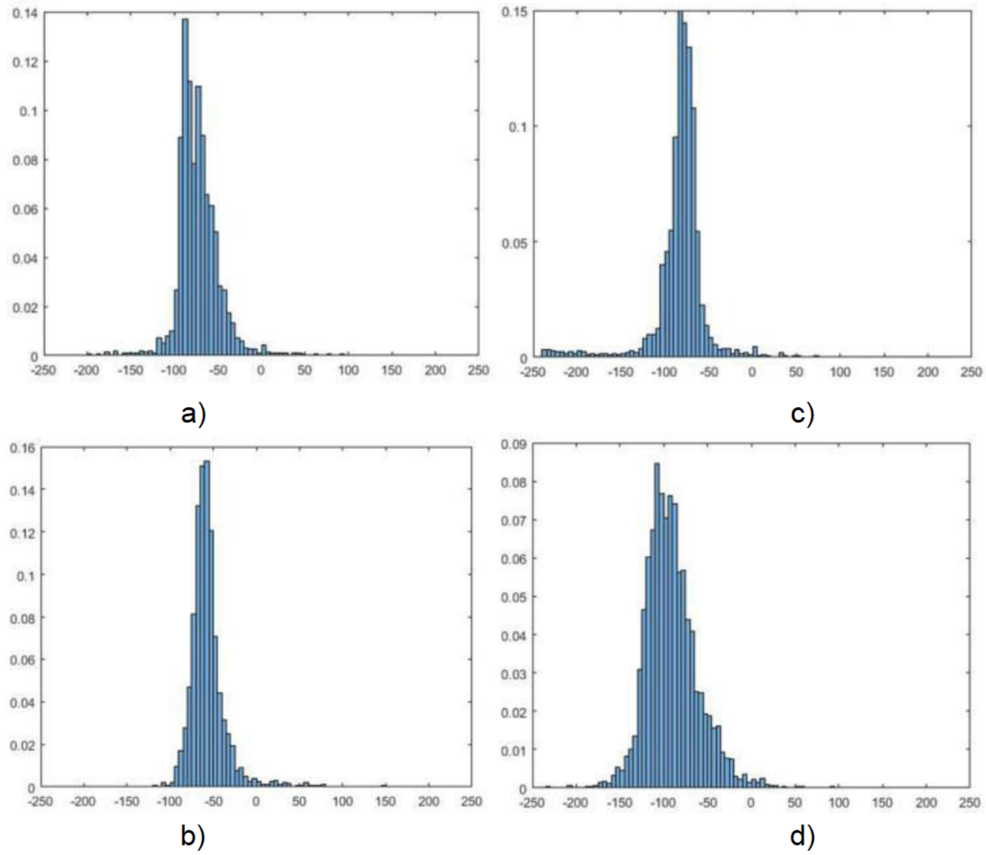


Figure 4: Histograms of the sheet perception with the text: **a**) “k”; **b**) An-2; white sheet: **c**) “k”; **d**) An-2.

The histograms (Figure **4**) show that when they work at a computer, a little-known process of perception spatial interaction occurs with a text and a white sheet. They are perceived between the eyes and the monitor screen. Similar processes occur when they observe the stereoscopic depth of various types of stereograms. We believe that it is the shift in the depth perception outside the plane of the two-dimensional image location of various types of stereograms that stimulate the development of the 3D phenomenon (Minzaripov *et al*., 2009). In other words, everyone who works behind the screen of a personal computer, regardless of his desire, converts planar perception into three-dimensional images. I.e. a 3D phenomenon development is mastered. The initial conditions for the formation of the three-dimensional perception of planar images is the effect of relief. Relief is the separation and depth of some planar images from others. The relief was revealed by us on some stimulus images for the sample of about 2000 young people (14-22 years).

CONCLUSIONS

1. The plot "I". The comparison of three types of stimulus images shows that visual perception of planar image perspective was revealed: 45% of students with "Kazan University" painting and 36% - "The Seine near Giverny". The researcher An-2 is among them.

2. The plot "II". We assume that when they work with a text on a personal computer monitor, the condition is fulfilled for the development of the 3D phenomenon initial elements. In other words, there is an ongoing training process for planar image conversion into three-dimensional projections.

3. Regarding the plane of the monitor screen, there is no unambiguous perception of the planar image perspective location.

4. The statistical application of the mathematical-graphical method is shown to detect the perception of depth and spatial perspective of planar images.

SUMMARY

Registration of eye bifixation points using a portable binocular eye-tracking allows you to explore the perception of depth, and the spatial volume of the planar image spatial perspective. The use of 3D raster image techniques, the construction of the difference histograms ΔX unambiguously show the possibilities of an objective method for spatial attribute record of 2D images.

Currently, only the first experimental results have been obtained. The surveys conducted among the respondents (19-21) clearly indicate that the vast majority already perceives the depth and volume of the used stimulus images.

We have no reliable assumptions about the reason for the 3D phenomenon development. We can agree with the Ringach hypothesis (Ringach *et al*., 1996) that it is possible that there is a single three-dimensional representation of surfaces where depth information from multiple visual cues is combined and that the estimated three-dimensional structure drives vergence eye movements. This hypothesis assumes that various visual cues affect the perceived three-dimensional shape and can serve as effective stimuli to vergence eye movements.

Nowadays, the visual perception of youth creates new conditions for planar image perception. There is a real perception of the depth, volume, spatial perspective of the paintings and other planar scenes.

It is possible that the perception of three-dimensional attributes of planar images can characterize the creative elements of thinking (Fazlyyyakhmatov & Antipov, 2019).

ACKNOWLEDGEMENTS

The work is performed according to the Russian Government Program of Competitive Growth of Kazan Federal University.

AUTHORS CONTRIBUTION

The authors prepared and conducted research on the article «eye movement features in the depth and spatial perspective perception of paintings and other static scenes»

Marsel Fazlyyyakhmatov, topic selection, study preparation, preparation of results.

Adelina Mutagirova, data collection, writing the main part.

Oleg Nedopekin, performed conceptualization, methodology.

Vladimir Antipov performed the conclusion and structuring of the text.

REFERENCES

Abromavicius, V., Serackis, A., Katkevicius, A., & Plonis, D. (2018). Evaluation of EEG-based Complementary Features for Assessment of Visual Discomfort based on Stable Depth Perception Time. *Radioengineering*, *27*(4), 1138-1146. https://doi.org/10.13164/re.2018.1138

Angelaki, D.E., & Hess, B.J. (2005). Self-motion-induced eye movements: effects on visual acuity and navigation. *Nature Reviews Neuroscience, 6,* 966–976. https://doi.org/10.17223/17267080/67/11

Antipov, V.N., & Fazlyyyakhmatov, M.G. (2018). Evaluating Model of Conditions for Forming Volumetric Visual Perception of Flat Images. *Sibirskiy Psikhologicheskiy Zhurnal-Siberian Journal of Psychology*, *67*, 149–171.

Antipov, V.N., & Zhegallo, A.V. (2014). Three-dimensional perception of planar images in a computerized environment. *Experimental Psychology, 7*(3), 97–111.

Antipov, V.N., Minzaripov, R.G., Fazlyyyakhmatov, M.G., Yakushev, R.S., & Zhegallo, A.V. (2019). On the Impact of Visual Information on Religion, Culture and Art", *Lecture Notes in Arts and Humanities*, *315*, 55–58.

Antipov, V.N., Zhegallo, A.V., Galimullin, D.Z., & Fazlyyyakhmatov, M.G. (2018). The experimental study of the visual perception of depth the flat images, eye movement registration. *Psychology, Journal of the Higher School of Economics, 15*(2), 384–399.

Backus, B.T., Fleet, D.J., Parker, A.J., Heeger, D.J. (2001). Human cortical activity correlates with stereoscopic depth perception. *Journal of Neurophysiology*, 86, 2054–2068. https://doi.org/10.1152/jn.2001.86.4.2054

Barabanschikov, V.A., & Zhegallo, A.V. (2013). *Registration and analysis of the human gaze direction. Moscow: Publishing House "Institute of Psychology RAS"*, 316 p.

Bucci, M.P., Bremond-Gignac, D., & Kapoula, Z.A.A. (2008). Poor binocular coordination of saccades in dyslexic children. *Graefe's Archive for Clinical and Experimental Ophthalmology*, *246*, 417-428. https://doi.org/10.1007/s00417-007-0723-1

Cazzato, D., Leo, M., Distante, C., & Voos, H. (2020). When i look into your eyes: A survey on computer vision contributions for human gaze estimation and tracking. Sensors, 20(13), 3739. https://doi.org/10.3390/s20133739

Fazlyyyakhmatov, M., & Antipov, V. (2019). Creative and Cognitive Perception of the Planar Images. *International Journal of Education Sciences*, *27*(1-3), 70-77. https://doi.org/10.31901/24566322.2019/27.1-3.1105

Fazlyyyakhmatov, M., Zwezdochkina, N., & Antipov, V. (2018). The EEG Activity during Binocular Depth Perception of 2D Images. *Computational Intelligence and Neuroscience*, 2018, Art. No. 5623165. https://doi.org/10.1155/2018/5623165

Gómez-Huélamo, C., Del Egido, J., Bergasa, L. M., Barea, R., Ocana, M., Arango, F., & Gútierrez, R. (2020). Real-time bird's eye view multi-object tracking system based on fast encoders for object detection. In 2020 IEEE Intelligent Transportation Systems Conference (ITSC). IEEE

Gregory, R.L. (1997). *Eye and brain: the psychology of seeing.* Princeton University Press. https://doi.org/10.1515/9781400866861

Iijima, A., Komagata, S., & Kiryu, T. (2012). Vergence Eye Movements Signifying 3D Depth Perception from 2D Movies", *Displays*, *33*(2), 91–97. https://doi.org/10.1016/j.displa.2011.11.001

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and futures words on fixation durations. *Journal of Experimental Psychology*, *135*, 12-35. https://doi.org/10.1037/0096-3445.135.1.12

Liu, M., & Nijhuis, S. (2020). Mapping landscape spaces: Methods for understanding spatial-visual characteristics in landscape design. Environmental Impact Assessment Review, 82, 106376. https://doi.org/10.1016/j.eiar.2020.106376

Minzaripov, R.G., Antipov, V.N., Chitalin, N.A., Shaposhnikov, D.A., Baltina, T.V., Skobeltsyna, E.G., & Yakushev, R.S. (2009). On the application of volumetric creative cognitive vision development methods in the innovative educational space. *Scientific notes of Kazan University. Series: natural sciences, 151*(3), 266-277.

Ringach, D.L., Hawken, M.J., & Shapley, R. (1996). Binocular eye movements caused by the perception of three-dimensional structure from motion. *Vision Research*, *36*, 1479–1492. https://doi.org/10.1016/0042-6989(95)00285-5

Schutz, A.C., Braun, D.I., & Gegenfurtner, K.R. (2011). Eye movements and perception: A selective review. *Journal of Vision*, *11*(5), 9, 1–30. https://doi.org/10.1167/11.5.9

Yao, Y., Liang, X., & Zhao, J. (2018). Binocular Depth Perception in Psychological and Clinical Studies. *Advances in Psychology*, *8*(12), 1795-1803. https://doi.org/10.12677/AP.2018.812209

Zhegallo, A., & Marmalyuk, P. (2015). ETRAN–R Extension. Package for Eye Tracking. *Results Analysis, Perception*, *44*(8–9), 1129–1135. https://doi.org/10.1177/0301006615594944

Received on 25-10-2020 Accepted on 27-11-2020 Published on 17-12-2020

DOI: https://doi.org/10.6000/1929-4409.2020.09.198